SOME PSYCHOLOGICAL ILLUSTRATIONS OF THE THEOREMS OF BERNOULLI AND POISSON.

By E. W. SCRIPTURE, Ph. D., Yale University.

Professor Bruns (*Ueber die Ausgleichung statistischer Zählungen in der Psychophysik*, *Phil. Stud.*, 1893, IX. p. 1) has pointed out the loss of energy in psychology consequent on the ignorance of the past development of the exact sciences. It is my desire in the present article to show how we can distinguish good psychological work from indifferent and bad by employing means known to every mathematician, but generally unknown to psychologists. The fact that I make use of recent publications by Professor Münsterberg for this purpose is not in any way to be taken as a disparagement of his work. The statements made in his "Studies from the Harvard Psychological Laboratory" in the first number of the *Psychological Review* for 1894, led me to inquire how far the conclusions drawn were justified by the facts given. As he has not verified his results by the proper mathematical means, the labor fell upon his readers.

In the chapter on Optical Time-Content, a series of conclusions is drawn in regard to the overestimation of an interval of time filled with one kind of visual impression as compared with an interval filled with another kind. The first question to ask is this: Accepting the facts exactly as stated, how much reliance is to be placed on the conclusions drawn from them? For all statistical work a definite answer was given by Jacob Bernoulli. Briefly stated, it is this: If there are n total cases in which A occurs r times and is absent s times, then the certainty that the observed result r agrees with the true result within the limits

$$l = \pm \gamma \sqrt{\frac{2rs}{n}}$$

is given by

$$P = \frac{2}{\sqrt{\pi}} \int_0^{\gamma} e^{-t^2} dt + \frac{\sqrt{n} \; e^{-\gamma^2}}{\sqrt{2 \pi r s}}$$

In case III. there are results given for six persons in their overestimations of a changing color as compared with a monotonous one in 100 experiments each. The arithmetical mean of all the results is fifty-nine, but the range of variation extends from fifty-one for J. to sixty-five for P. The average range of variation is thus, $l = 7$. With what confidence can we state that the actual average fifty-nine expresses the truth? Substituting the values, $l = 7$, $r = 59$, $s = 41$, $n = 6$, in the above equations, we find that $P = 24\ \%$. Or, in other words, we would be justified in expecting that once out of every

four times a result lying somewhere between fifty-one and sixty-five would be obtained, which—it must be confessed—is a very small amount of trustworthiness in scientific work.

What should Professor Münsterberg have done? The answer is given by a reversal of Bayes's theorem, which I believe I am justified in making. It can be stated as follows: If the probability of the occurrence of an event is p, and of its non-occurrence is q, then, in order to assign a limit, l, with a certainty of

$$P = \frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-t^2} dt$$

the total number of cases must be $n = \frac{2\gamma^2 p q}{l^2}$. As the experiments were made in hundreds, the numbers fifty-one, fifty-nine and sixty-five are also percentages. Thus $p = 59\%$ and $q = 41\%$. In statistical work we generally demand a degree of certainty.

$$P = 0.999978$$

on a scale of 1 = infinite or absolute certainty. For this value, $\gamma = 3$. By carrying out the processes indicated, we find that $n = 872$. From which we must conclude that with so inaccurate a method and with such variation among the individuals, a conclusion is justifiable only on the basis of experiments on nearly 900 individuals. Every statistician knows that experiments on a few individuals are of no value unless there is practically no variation among them. On the basis of six individuals, Professor Münsterberg says: "It appears to me unquestionably that, etc." Case III. is, moreover, flatly contradicted by case II., where the individuals disagree completely.

A very pretty illustration of one of Poisson's theorems can be drawn from Professor Münsterberg's statements on "Memory," p. 36. Two series, each of 2,140 presentations, visible and audible, were compared as to the frequency of mistakes. How large must be the difference between the results, in order to justify the conclusion that they are really different? The theorem can be stated thus: If in two groups, n and n^1, the observed frequencies are p, q, p^1, q^1, then we can conclude with a certainty of

$$P = \frac{2}{\sqrt{\pi}} \int_0^\gamma e^{-t^2} dt$$

that the difference between the true probabilities R, R^1, will be

$$K = R - R^1 = p - p^1 \pm \gamma \sqrt{\frac{2\,p\,q}{n} + \frac{2\,p^1\,q^1}{n^1}}$$

Let us take the most unfavorable case, W., which gives 25.1 % for the visual series and 31.6 % for the auditory. Here $n = n^1 = 2{,}140$, $p = 0.251$, $q = 0.749$, $p^1 = 0.316$, $q^1 = 0.684$, and, with a demanded certainty of $P = 0.999978$, $\gamma = 3$. The computation gives 0.058 for the limits within which the observed difference may vary from the true difference. The observed difference is $31.6 - 25.1 = 0.065$, which exceeds the limits of possible variation. Professor Münsterberg is thus justified in concluding that the visual memory excels the auditory memory under the given circumstances.